

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAACAGAAAACCTGTTAGAAATGTTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCTTGTACTTGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCTT
CCAGAAAAACAACCTTTTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCTTCTTT
GGTTTTTCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFI~~FSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV~~
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLT~~YIRD~~FQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTTCTTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTTTG
GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAG
GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGG
TTTTCATGGTGCCCTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAAACGA
CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTC
CATTCTCAGCCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTTTCTGGATTGGTGTCTCAACTGCCCATACACTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
CATAAGCAAAAAGAAAAGGATGGCAATGGCAGGAGAACAAATGTTCCAGAAGGGGGAAGTGCATAACA
AACCATCAGGTTTCTGGGGAATGATAAAAAGTGTACCCTTCAGCATCAGGAAGTGAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTTAATTTTCTTGGTT
ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTT
GGGAAAACGGATCCTGTCACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT
GAAGTTTGGTCCCAACACATTTCTTCATTCTGTTGGAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCTTGTCTG
CTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCCTT
AGAATACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTG
ATGTGATCTTCCCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTA
GATATAAGAGGGGGGAAAAATGGAACCGAGGCTGACATTTTATAAACAAACAAAATGCTATGGTAGC
ATTTTTACCTTCATAGCATACTCCTTCCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGGAGAACTAATCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
TGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTACATGGAAAAGGTTATAGCTTTG
CCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYPFHWMNLCVILLILVFMVPPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGD
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDAL EELS
RQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTD
PVTGRIETVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIM
GMYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAP
EQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT
 GTCATGCAGAATTATGGGGATCACCTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG
 AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTTGGAGATGGATTTCGTGGTCATCTCTAG
 GATTAGCCCAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCTGATTGGAAGGTTCCAGTGA
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAAGTGCATTCCAGAA
 ATTATCACCACCAAGATCCCATATTCAACACTCAAAGTCAACACAAACAGAAATTTATTGT
 CAGTGACAGTACCTACTCGGTGGCATCCCCCTACTCTACAATACCTGCCCCCTACTACTCCTC
 CTGCTCCAGCTTCCACTTCTATTCACGGAGAAAAAATTGATTGTGTACAGAAGTTTTATG
 GAAACTAGCACCATTGTCTACAGAACTGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGA
 AGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGTCTCCTCTTCTTTGGTGCTG
 CAGCTGGTCTTGGATTTTGTATGTCAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAAT
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAA
 TGAGGAATCAAAGAAACTGATAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGC
 GATGCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTT
 CATGCTCCTTACCCTGCCAGCTGGGAAATCAAAGGGCCAAAGAACCAGAAAGAAAGTCCA
 CCCTTGGTTCCCTAAGTGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAAT
 GCCCTTCTCCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT
 TTCTAGCCTGGCTATGTCTAATAATATCCCAGTGGGAGAAAGGAGTTTTGCAAAGTGAAGGAC
 CTAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTG
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGAGCTCAGACCCTTCTTCA
 GCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCTTCTGAGCCCGTAAGAGCAAAAGAAT
 GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATACTTGAACCTAATCTCTGTAAA
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTGAGCAGGGACTGTAAAC
 ACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA
 AAAATAAAAACCTCTTATAAATTTCTATTTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATCAACAAACATTGCTGAATAGCTACTATATGTC
 AAGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAATTGCACATAGTAG
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT
 TTTTATTTTTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATT
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTTTAAAGTGCC
 ATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGACTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAAGCATTAGAAAACTT

FIGURE 6

MARCFSLVLLLSIWTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKA AFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFFFTNKNQQKEMIETKVVKEEKANDSNPNESKKT DKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCGCCGCCACCGCGCCGCTCCCGCATCTGCACCCGAGCCCGGC
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCGGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATGCGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG
CGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCCG
GCTCTCAGCTACCCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAACTGAT
GGAGGACACGCAGCACAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACAGAC
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGTGTGGAGACCA
GCTGTGTGTCTGGGCTCACTGCACCAAATGGCCACCAGGGGCAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGTGTTCCCTGTGTGC
ACACCCCTGCCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGAAGCCGACCTTCGTGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTCTCTA
CATCTTCTTCCAGTAAGTTTCCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTCAGCT
CCCCCAGGCTGTTCTCCAGGCTTACAGTCTGGTGTGCTTGGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCTCTACAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCTGCTTTGCAAACATCAACCTGGCAAAATG
CAACAAATGAATTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTGTC
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT
CTCCACTACCCACACCAAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAGGAAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACCTAATCTCACATCCCTCTAAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGCAGCCGTCCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATG
TTTTCAGGTGTGATGGACTGTTGCCACCATGTATTATCCAGAGTTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAA
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAAKASSEVNLANLPSPYHNETNTDTKVGNNTIHVHREIHKITNNQTGQMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQMLCTRDSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPGGCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQLEDLE
RSLTEEMALGEPAAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCCTT
AACTGGGTACTGGCCCTGGGCCAATGCGTCCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTCAATTGGCATTGAGCCCTCATCCTGACCCTTGTGCAGATAGCCCGGGTCATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGTGCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTGAGGGTGGTCTGCTGGACAAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
GGTGGGAGGCGTGGGGCTCTGTCTTCTTTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGGCCTAT
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTCCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
AGATTCTGGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCACCCTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
AAACCTCCGTCTCTATTAATAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTATCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTGCAGTGAGCCGA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAA
AAGATTTTATTAAAGATATTTTGTTAACCTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVLDKVTDLLLFFGKLLVGGVGVLSEFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCCGGGCGCCGAAGCCGGGAGCCACCGCCCATGGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCGC
CAGCCGAACTCCACCGTGAGCCGCTCATCTTCACGTTCTTCCTCTTCCTGGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCCGGCGTGGAGAGTCACTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC
GGGATCCCCACCGTCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTCTTTTACCCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCCGGGTGCCATCCAGAATGGGTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGTCCCTTCACCAACATCTGGTTCTACTTCGGCGTCTGCG
CTCCTTCCTCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTTGCGCACTCCTGGAACAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCATTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTCACTCTCCTC
TTCTACTTGCTGTGATCGCGCCGTGGCGCTGATGTTTATGTACTACACTGAGCCAGCGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGTCTCTGC
CCAAGGTCCAGGACGCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTTCATCACCTCTACACCATG
TTTGTACCTGGTCAAGCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCCGAGCATTG
TGGGCCTCATCATCTTCCTCTGTGCACCTCTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGT
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGCT
GGCAGCCTGTGAGGGCCGGGCCTTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCC
ACTTCTGCCTGGTGTGCGCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT
CCTCTACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTGCCCCTGAGCCGGGC
CTTCTAGTCGTAGTGCCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCTCTTCCTTCCCTCCTCCTGTTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACCTCCACCCAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCAGGGGACCCTGCCCCCTTCTCTG
GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAAACAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLLSASCCLCGSAPCILSCCPASRNSTVSRLIFTFFFLGLVLSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFILIQLVLLIDFAHSWNQRWLKAE
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVCSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAAACEGRAFDNEQDGVITYSY
SFFHFCLVLASLHVMTLTNWKPGETRMISTWTAVVVKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCTCCTCCAGGATGAACCACTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTTGTGTGTACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG
TGATCCTTTTCAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTTCATTCATC
CTTGCCCTGGATTGAGACGTGGTTCCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
GTCAGTTTTATTCCCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCACCTTTTAGAAGTATGAGTACTACTTTTGTTAAATGTGAAAAACCCTCACAGAAAGTC
ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCCCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTTTAGTAAGCAAGATACCTTTTTTATTTCATTCAC
AGAATGGAAATTTTTTGTTCATGTCTCAGATTTATTTGTATTCTTTTTTAACACTCTACATT
TCCCTTGTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGAATAAAATCTG
ACATGTCAATGTGGCTAGTTTTATTTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAATGTCACCAGACATTTGTATTATTT
TTATCATGAAATCATGTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC
ACAAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCGACCCGCCAGGAAAGACTGAGG
CCGCGGCCTGCCCCGCGCGGCTCCCTGCGCCGCGCGCCCTCCCGGGACAGAAGATGCTGCTCCAG
GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGGTGACGGGCTGCCCCAT
CCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCCGCCAGGGGACCACGGTGCCCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
AGGCAGCTTTGCCGGCCTGCCGGGCTGCAGTCTCTGGACCTGTCACAGAACCAGATCGCCAGCC
TGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCGGGCACTGCCCCGCTGCGCCTGCCCCGCTGCTGCTGCTGGACCTCAGC
CACAAACGCCTCTTGGCCCTGGAGCCCGGCATCTTGGACACTGCCAACCTGGAGGCGCTGCGGCT
GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTACCGCCTTGGCAACCTCCACGACC
TGGATGTGTCCGACAACAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCTGACG
CGCCTGCGGCTGGCCGGCAACACCCGCTTGGCCAGCTGCGGCCCGAGGACCTGGCCGGCTGGC
TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCGCTGCTGGCGACCTCTCGGGCC
TCTTCCCCCGCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC
TGGTTTGGCCCCCTGGGTGCGCGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCA
CTTCCCCCACAAGACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCCCAG
CCACCACCACACAGCCACAGTGCCCAACACGAGGCGCGTGGTGCGGGAGCCACAGCCTTGTCT
TCTAGCTTGGCTCTTACCTGGCTTAGCCCCACAGCGCGGCCACTGAGGCCCGCCGCGCCCTC
CACTGCCCCACCGACTGTAGGGCCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTCGCTCA
ATGGGGGCACATGCCACCTGGGGACACGGCACCTGGCGTGCTTGTGCCCCGAAGGCTTACG
GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCACGCCGAG
GCCACCACGGTCCCTGACCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
GGCCCTGATAAGCGGTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACAGGTCACCCA
GCTGCGGGCCCAACGCCACTTACTCCGTCTGTGTGTCATGCCTTTGGGGCCCGGGCGGGTGCCGGAGG
GCGAGGAGGCCTGCGGGGAGGCCCATACACCCCGAGCCGTCCACTCCAACCACGCCCCAGTCACC
CAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC
GCTGGCTGCGGTGGGGGACGCTACTGTGTGCGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG
ACAAAGGCGAGGTGGGGCCAGGGGCTGGGGCCCCGGAAGTGGAGGGAGTGAAGGTCCCCCTGGAG
CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
ACTCATGGGCTTCCAGGGCCTGGCCTCCAGTCAACCCCTCCAGCAAGGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCTGCTGCC
ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGACAGAGGGCTGTGTGACCACAGCT
GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
CTAACGTCCCCAGAACCGAGTGCCATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCCTGGGTCTGCTGGGCTCTCCAC
TCCAGGCGGACCCTGGGGGCCAGTGAAGGAAGCTCCCGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCAAGGAACAAAAGAACTGGAAAGGAAGATGC
TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTCCCATTATTCT
GGGAAGATGTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAAAGACAAACGATGATATGAA
GGCCTTTTGTAAAGAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLPLLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTPRDPVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILD TANVE
ALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSLQALPGDLSGLFPRLRLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPVTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPDCCPSTCLNGGTCHLGRHHLACLCEGFTGLYCESQMGQTRPSPTP
VTPRPPRSITLGIIEPVSPSTSLRVGLQRYLQGS SVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCMFLGPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAYCVRGRGRAMAAQAQDKGVGPGAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFPGLQSPHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTAGAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGCACAG
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGGAATGAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGCTCTTGG
CTTTCTGTATGCCTCTGGACTTGGTGTAAATCAAGTCAGGCAAAGGCTCTGTATATTATACAT
TTGGAGCTCTTGGGGGCAATCTAATAGCCACATGGTTTTGGTAAGTAGACTTTAGTGGGAAGGCT
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTTCAGCTTTTCATGATC
CAGATTTGCTTGTATTAAGACCAAATATTCAGTTGAACCTCCTTCAAATCTTGTTAATGGATAT
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTTCTTTAAATGATTAG
TTTGGCTGATTGCCCCTAAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG
GAATTGTCAGAATCATTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA
AAATTTTGTAAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTTGCAACAATGC
CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTTCCATTATTCTGGTCATTTATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTCTTTGCTAATTTGG
AAGATTAACCTATTTTAAATAAAATTATGTCTAAGATTAAAAA
AAA

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTLTSDSVKDHTTAGRVVAGQIFLDSESEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFPFLDK
EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAARRMQEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFELTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLIAMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAAGCCCATTTCTGCAGTGGAATTTTCATGAACTAGCAAGAGGACACCATCTTCTT
GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGAAAAATGCTCTTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAATTTTGAT
AAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGT
CCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTGCGCTTGAATCGTTGGAGGGGGCTATACTCCATCCAAA
TATGCAGTGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAA
AGTCTAGACAAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTTAGGCTTTGCCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCAGCATTTACAGTAACTTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTTAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKG FHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNOVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGNKSYVNMD
LSPVVECMDHALTSLEPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
CATTTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATCCGT
CGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACCTTGCAAGGAGCATT
TTCAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
ATCGACTGGAACATTCCCTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGGTT
GCCAATCTGGGCATGTCTGAACAACCTGGGTTATAAACTGTATCAGGTTCCGTGTATGTCCACTGG
TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTTGAAGAAGATGGATCCTTAAAGGAGG
TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA
GAAAAGGAGAGGAGCAGATTTCAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
ACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTCTGTGTT
ATGTCTTTAAAAAATAGACATGTTTCTAAAGTAGCTGTAACCTACAACCACCATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
CTACATTTTGATCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC
ATTTCTATTTGTTTTTACTATGTTGAGCTACTTGCAAGTTCATTTGTTTTTACTATGTTTACC
TGTTTGAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA
GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
ACACATTTCTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
GAGACCATGTCTATTAATAAATAAATGGAAGCAAGAATAGCCTATTTTCAAATATGGAAA
GAAATTTATATGAAAATTTATCTGAGTCATTAAATTTCTCCTTAAGTGATACTTTTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
AAAATTTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLGGEVKGEAKNSITDSQMDDVEVVYTIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSSRAV
QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHSVSKSSCNYNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLDQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCCGCGCCGCCACACCCTCTGCGGTCCCCGCGGCGCTGCCACCCCTCCCTCCTTCCCC
GCGTCCCCGCGCTCGCCGGGCGAGTCAGCTTGCCGGGTTGCGTGCCCCGCGAAACCCGAGGTCACCAGCCGCGCTCT
GCTTCCCTGGGCGCGCGCGCTCCACGCCCTCCTTCTCCCTGGCCGCGCGCTGGCACCGGGGACCGTTGCGCTGA
CGCGAGGCGCAGCTCTACTTTTGGCCCGCGTCTCCTCGCGCTGCTCGCCTCTTCCACCACTCCAACCTCCTTCTCCC
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGCGCCGCTTCCGCTCCGGTCCCAA
GGTGGGAACGCGTCCGCCCGCGCCGACCATGGCACGCTTCGGCTTGCCCGCGCTTCTCTGCACCCTGGCAGTGCTC
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGACGCTTTTACGTGTCCAAAGGCTTC
AACAAGACGATGCCCCCTCCACGAGATCAACGGTGATCATTGAAGATCTGTCCCAGGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAAAAGTGTGGTCAGCGAACAGTGCAATCATTG
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTATGAATTTCTCAAAGAACTACTTGAAATGCAGAGAAATCCCTG
AATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAAATTTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTG
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCTGGAGCGGATG
TTCCGCGCTGGTGAACCTCCAGTACCACTTTACAGATGACTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTCGGAGATGCTCCTCGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTGCTCAAGGC
TTAGCGGTTGGGGAGATGTCGTGAGCAAGGTTCCGTTGTAACCCACAGCCAGTGATCCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGTCTCGTGACTGTGAAGCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCTTTCAACATTGAATCGGTGATGGATCCCATGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTTCAGTGTCTCAGAAGGTTTTCAGGGATGTGGACCCCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCCCTTCAGTGCTCGCTTACAGCCACATCACCCGAGGAACGCCCAACACAGCAGCTGGCACT
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATCTGGTCTCCCTTCCGAGCAAC
GTTTGCACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAGGCAAAAGCAGGTAC
CTGTTTGCAGTGACAGGAAATGGATTAGCCAACCAAGGCAACAACCCAGAGGTCAGGTTGACACCAGCAAACAGAC
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGCTCGGGGCA
CAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCTGTTATGCAGAGAGTGGAGATTAATTTCTCAAACCTCTGAG
AAAAAGTGTTCATCAAAAAGTTAAAGGCACCAAGTTATCACTTTTCTACCATCCTAGTGACTTTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAGTGTGACTTTGTTTTCTCATTGAGTTTGGG
AGGAAAAGGACTGTGCATTGAGTTGGTTCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGAGGTACAGAA
CTATAGTTAGTTGTGATTTGTGATTTTATCACTCTATTATTGTTTGTATGTTTTTTCTCATTTCGTTTGTGGGT
TTTTTTTCCAACGTGATCTCGCCTTGTTTCTTACAAGCAAACAGGGTCCCTTCTTGGCACGTAACATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTATCATGTTATCTTATTAAGAAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKS LNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVG NVNLEEMLNDFWARLLERMFRLVNSQYHFTDEYLECVSKYTE
QLKPFQGDVPRKLLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL
VTVKPCYNYCSNIMRGCLANQGDLDFEWN NFIDAML MVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKV FQGC GPPKPLPAGRISR SISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNPEVQVDT S
KPDILILRQIMALRVMTSKMK NAYNGNDVDFFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCCACTACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTTCAAGATCATTTTGTGTTGCTCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT
AGCTAGTGTCAATTTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISSLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGTGCTTCTTACCCCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
CAACTTTCAGTTCTACCCACCGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCTTTGAGA
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTGCTG
CTCTGTGCAGAGCCCAAGGAAGGTCCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTTC
GAGCCCACCTGGAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
GAACGCCCAGTTCTCCGAAATCCAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTGCTCC
TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTA
GCAGAATGAGAGAAGACATTGATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCTC
CCAATGTTGTCCCTTTCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCTCCTGAGGCTACAC
CCATGCGTCTCTAGGAAGTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTLPPLHMLLGCWQPLCKSYFPYLMAVLTPKSNRMESKKRELFQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQVFPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFSSKALICSFPSL
QLEQATHQPIYLPRLGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGGCCAGAAAGTCTCTTCTG
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTATTAAAGAAATT
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

APP ID=10063569

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAAATCATGTCGG
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC
TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGTTATTTTGGGATTGTTGTT
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTTCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGGAATATAAGCCCCTTTCGGGCATTCCGTACATGTGGTCGTACCATT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAG
TGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC
TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAG
GATTCCGAGAATCATTGTCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
AACCAGAATGCATATACTACAACGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
ATTCAAAATCTTGTCGAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA
TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCTGTGTTTTGCTGTTGATC
TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTCTGAGTTTCGTA
AAAAGGAGCAACAAATTAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT
TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTT
TTAAAGACCTAATAAACCTATTCTTCTCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTTLLVHIFISLVILGLLFVCGVLWWLYDYDNDLSIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA
GAVVTCYFNRSKNDDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIIFLGKVLVVCFTVFGGLMAFNYNRAQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCTTAGA
ATAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCATTCTGTGTGGT
GAAAATTTTTTGAAGGAAATTTGCCTTCTTCAAACAAGGGTGTCTTCTGATATTTATGAGGAC
TGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTTCTTGTGTTTGTGCTGGTGAAGTAC
ATTCAAACAAAGAAACGGCAAAGAAGATTAAAGGCCCAAGTCACTGTGCCTCAGATCAACTGC
GATGTCAAAGCCGGAAAGATCATCGATCCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG
TACACAGTGGTGTGCTTGATAATTCAGGAGGAAAAATACTTGTTCGGAAGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
TATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAACTACCCATCAGCTCTTACATACTCATCAT
CGAAAAGTCCAGCTGCCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACCTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
CACCTTGCCAAGGCCATCCCCTTCTGTCTTCTACCACCAGCATCCCAGACCACAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC
AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGC
GGATGTCAAGCCTGGGACTTGTTCAAAAGAAGAAATGAGCACACAGTCTTTGGAGCCAGTATCCC
TGGGAGATCCAAATGCAAAATTGACTTGTCTGTTTTAATTGATGGGAGCACCAGCATTGGCAAA
CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCCAAGCTCTTGACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCCTGCTACTCACTTTAACCTCAAGACAC
ACACGAATCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT
GTAGGTGCGGCCATCTCCTTTGTGACCAAGAACTTCTTTTCCAAAGCCAATGGAAACAGAAGCGG
GGCTCCCAATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
GACTTGGGAGAGAGTCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGCCCAACTTTGCAACAAAGGCCGTGTCAGAACAAACGGCTTCTACTC
GCTCCACGTGCAGAGCTGGTTTGGCCTCCACAAGACCCTGCAGCCTCTGGTGAAGCGGGTCTGCG
ACACTGACCGCCTGGCCTGCAGCAAGACCTGCTTGAACCTCGGCTGACATTGGCTTCGTATCGAC
GGCTCCAGCAGTGTGGGGACGGGCACTTCCGCACCGTCTCCAGTTTGTGACCAACCTCACAA
AGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTGCAGTACACCTACGAACAGCGGC
TGGAGTTTGGGTTTCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGTGGGC
TACTGGAGTGGTGGCACCAGCACGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTTCAAGAA
GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATACCGACGGGAGGTCTACGACGAGCTCC
GGATCCCAGCCATGGCTGCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTTGCCTGGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCACTCCTTCTTTGTGGACGA
GTTTGACAACCTCCATCAGTATGTCCCAAGGATCATCCAGAACATTTGTACAGAGTTCAACTCAC
AGCCTCGGAAC**TGA**ATTCAAGCAGGCAGAGCACCAGCAAGTGCTGCTTTACTAACTGACGTGTT
GGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC
AAATGTCTTGTTATTATTCTTTGCCATCATGCTTTTTCATATTCCAAAACCTGGAGTTACAAAGA
TGATCACAAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGGAGATTTTACAT
TTTGACAATTGTTTTCAAATAAATGTTCCGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
AGCTTTTGTGAGATTTTAAAGTTGTTATTTCTGATTTGAACCTCTGTAACCCTCAGCAAGTTTCAT
TTTTGTGATGACAATGTGAATTTGCTGAATTAATGTTTAGAAGGATGAAAAATAAAAAAAA
AA
AAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKTHNTSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTQLPLVKRVCDDRLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAAWAAQEELEVIATHPARHSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTTTCTGTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATCCACAGTTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTTCCCCTGGCTGCTCTATCACGTGGTGTCTCCGACTACTCACCCCGAGTGTA
AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTAC
CTCAGCCTTCCCCACTACAATGTGATAGAACGCGTGAACGGATGTACTTCTATGAGTATGAGCCGATTTA
CAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTTCTGGTCATT
TGTTGACCTCCCACCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT
TGTTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAATGTGGC
ATTGTCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA
ACCTGACCTTGAAAACCATATGGCATTCAAGTGGGTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
AAGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
GAAGTTTTTCACAGGTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCATATTT
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTGGTGGTGGTGGTGGTGGT
TTGGTGCCAGGATCTATGAAATGATGGGTACGTAACCCATCAAGTTTGAAGATGTTTATGTCTGGGAT
CTGTTTGAAATTTATTAAGTGAACATTCAATTCAGAGACACAAATCTTTCTTCTATATAGAATCC
ATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCAAGGAGATCATCACTTTT
TGGCAGGTGATGCTAAGGAACACCACATGCCATTATTAAGTTCACATTCTACAAAAAGCCTAGAAGGACAG
GATACCTTGTGAAAGTGTTAAATAAGTAGGTACTGTGAAAATTCATGGGGAGGTCAAGTGTCTGGCTT
ACACTGAACTGAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAATATATAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGG
ACCAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGTAAAAACAAACATGTAGAGTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA
TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTCACTGAAGTTATA
CTGAACAAAATTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT
ATTATTTAAATTAAGTCAACTTTGTGTTTTTAAATGTTTGGAGATTTCAATACAAGATAAAAGGATAG
TGAATCATTCCTTACATGCAACATTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTTAAATATTTTACTGTGGT
AATATAGAGAAGATTAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQAEKEDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTCLKTMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQYEPFKVFPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLKVNHIHPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGGCAAGGATCGAGCTGGGTTGCTGGACGAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC
TCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGGAGTGCTGTG
AATATGATCAGATTGAGTGCCTCTGCCCCGAAAGAGGGAAGTCGTTGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCTGATCCACCCAGGTTGTACCATCTTTGAAAACGCAAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTTGTGGAAAGCTATCCCTTAAATGCTCACTGTGAATGGACCATTTCATGC
TAAACCTGGGTTTGTATCCAATAAGATTGTCTATGTTGAGTCTGGAGTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTTCGTGATGGAGACAACCGCATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGC
ATGCTCCTCATCCCTTGTTTCCATGACGGCAGCTGCGTCTTGAAGAAGCTGGATCTTACAAGTGTGCTGCTTGGCAGGC
TATACTGGGCGAGCTGTGAAAATCTCCTTGAAGAAAGAACTGCTCAGACCTGGGGGGCCAGTCAATGGGTACCAGAAAA
TAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATGGGCACCGTGGTGTCTTTCTTTTGAACAACTCCTATGT
TCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCCTGCCGA
GAACCAAGATTTAGACCTGGTGAGAGGAGAGTCTTCCGATGCAGGTTTCACTCAAGGGAGACACCATTACACAGCTAT
ACTCAGCGGCTTTCAGCAAGCAGAACTGCAGAGTGCCCTTACCAAGAAGCAGCCCTTCCCTTTGGAGATCTGCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCTTCTACCGCGCCTGGGCGAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGATGGGCGGGCACCATCTGCATCCCTATCTGCGGGAAAATGAGAACATCACTGCTCCAAAGA
CCCAGGGTTGCGCTGGCCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCACTGTGTTACTGACCTGGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTGGGGAAATTTACCGGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATTCTGCATCCCACTATGACCCCATCTGCTTGATGCTGACATCGCCATCCTGAAGCT
CCTAGACAAGGCCCGTATCAGCACCAGTCCAGCCCATCTGCCTCGCTGCCAGTCCGGATCTCAGCACTTCTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGTCTGCGCAGAGTGAAGGAGCCCTGGCTTCAAGAAGACACACTGCGCTCTGGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGTCCTTCCCGGA
CGAGCATCTCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCTTTCACCAAGGTGCTGCTTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTGGCCTGTGAACCTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAAGTACCTCCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGTCTTCCCAACTTTCAGTTATACGAATGCCATCAGCTTGACCGGGAAGATCTGGGCTTCATGAG
GCCCTTTTGAGGCTCTCAAGTCTAGAGAGCTGCCTGTGGGACAGCCAGGGCAGCAGCTGGGATGTGGTGCATGCCTT
TGTGTACATGGCCACAGTACAGTCTGGTCTTTTCTTCCCATCTCTTGTACACATTTTAAATAAATAAGGGTTGGCTTCT
GAAC TACAAA
AAA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFVVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGOII
KRVCGNERPAPIQSIGSSLHVLFSHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPQVQSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIENITAP
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVVSVDSSLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMLVSWSYDKTCSHRLSTAFTKVLFPKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATTTCTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTTGAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAAATGAAGGCGCTTCTGTTGC
TGGTCTTGCCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTGTATTCAGAACTCTGTA
AAGGTGCCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAAGATGGCTGTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCCTAGACA
ACCTTGCTACGTGTCTCGGCGAGGAGCGGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAACTA
GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTTCTTCGAA
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAATTCTGAAAACACCACTGCCCTGAAG
TCTTTCCAAGGTTGTACCACCTGATTCAGATGGTGAAATTACCAGCATCAAGATCAATCGAGTAGATCCAGTGAAA
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTATCGTGATGGGG
TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACTACGCTGTGCGTCTCCTGCGGAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAAGTCCGCA
GCAGGAACAATGGACAGGCCCCGGATGCCACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTTGAATAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTCATCTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA
GCCCAGAAAGTGGGCTCATCTGATTAGGCCAGTGAAAGACGTGTTACCTCGTGTGTCGCGCCAGGTTCCGGCAGC
GGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCAGGGCCAGGGGAGAGGAGCAACA
CTCCCAAGCCCTCCATCCTACAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCG
GCATGACCGTCGCGAGGGGAGCATCACATAGAGAATGGGATTGCTTATCTATGTCATCAGTGTGAGCCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAAGTGAAGAGTATGAGC
GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCCT
GGGTGATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTC
TGGGCTTCTGCATTGTAGGAGGTATGAGAATAACAATGGAACAAACCTTTTTTCATCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTGCTGTCAATGGTAGAAGTACATCAGGAATGA
TACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT
TTTTATAGAATCAATGATGGGTGAGGAAAACAGAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTATC
TTGTGAGTTTTTATATTAAAGAAAGAAATACATTGTAAGAAATGTGAGGAAAAGTATGATCATCTAATGAAAGCCAGTT
ACACCTCAGAAAATATGATTCAAAAAAATTAAGTACTAGTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC
AACATTGTTTATATTTTTCTATTCAATAAAAAAGCCCTAAACAATAAATGATTGATTGTATACCCCACTGAATT
CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCT
AAAATATTTTTTAAATGCATTGCTGAGAACGTTGCTTTCATCAAACAAGAATAAATTTTTTCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLS PANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGC PDGCASLTATAPS
PEVSAAATISLMTDEPLDNPAYVSSAEDGQPAISPVD SGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITS IKINRVDPSESLSIRLV
GGSETPLVHII IQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQ LGIKLVRKVDEPGVFIFNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESA AHLIQASERRVHLVVS RQVRQRSPDIFQEAGWNSNGSWSPG
PGRSNTPKPLHPTITCHEKV VNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
IKTG DILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGAAAAAGCTCAACTTGAAGCTTT
CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT
TTCCACCTTTCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTGGGTGGGC
CACCAGTAACACTTTCGTGGGTGCCATTCAAGAGATTCTTAAAGCAAAGGAGTTCATGGCTAATTTCC
ATAAGACCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
GACAACTGTCCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC
TTTGGGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG
CTTTACAGAGGGTCGCCATCTCTGTTCCCCACCGGAACAGAGAGAAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTTCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
TAAAAAGTTTAAATCGAGCCAAACTCTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTTGGG
ACTGCTTTATATCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
CATCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTG
GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACACTACTGGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTCCAGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT
TAGTATCTGTGGAACACAATCCTTTTATATATCAACATCACAGTGGATTTCTGGTTTGGTGCAATGACCC
TGGATCTTTTGGTGATGTTTGAAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTCTTTTGTATTTTCT
TAGCAGAGCTCCTGGTGATGTAGAGTATAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATATATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATACAATACTGTTATTCTTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTCAGGT
GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTCCGGTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCTTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT
TTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
ATTAATAATAATAATATGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAGCAAAA

FIGURE 42

MGFNLTFFHLSYKFRLLLLLTCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEENWDCFIHFDV
DLVPENDFNLYKCEEHPKHLVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNNYWGWWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCAACAG
ACGGGACAACCTGCAGAGCTGCAACCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCCGTCCCCTCCC
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHFFI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTTGTCAACATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAAGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCTGTGAGCAGAAACT
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCCATG
GTCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGG
AAACTCCTAACATATGCCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCCTCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIIYLWQLTGSAAAGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIIVTQNRNRERVDFFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCCMEHGEEDEVITWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRRERQEYIE
EKRRVDICRETPNICPHSGENTHEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
CTTGCACTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTATAGGTCTATTGCTTGTGGAATTCTGGAGGTCCTGTTGGGCT
CAGTCAGATAGTCATCGGTTTCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCM LISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHF DSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCCT
TTCTGATCAGGAGGCTTCTTTATGAATTAAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAGIAAVLSGKCKYKS
SQKQHSVPVEKAIPILITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCATTCGACACGGAGCAGATGCTGTCCGGGCTCCTGGCAGGGG
GTGCCTGGCCACAGTGGTGTCTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG
GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCCAGCCTGGCTATGGTTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
CATAAACCCGGGTGTGAAAAGCCAGGGAATGAAGCCCGGGGAGCGGGGAATCTGGGATTCAGGG
CTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTG
GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCTGGAA
GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
GCATCCCGTGAACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAA
AACACCACCTCTCATCTACTAATCTCAGCCCTTGCCCTTGAATAAACCTTAGCTGCCCCACAAA
AA
AA

FIGURE 52

MKFQGPLACLLALLCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVF
GHSRAWETSGGHGIFGSQGLGGQGGNPGGLGTPWVHGYPGNSAGSFGMNPQCAPWQGGNGGPPNF
GTNTQCAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGSGSQSGSSGSGSNGDNNGSSSGGS
SSGSSSGSSSGSGSGSGSGSGSGSGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPGCEKPGNE
ARGSGESGIQGRGQGVSSNMREISKEGNRLGGSGDNIRGQSSWGSGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCTGGCTTGGACCTATGCCCTTCTATAACAAGTCCCGCCGGCTCCAGTGTTTCC
CACAGCCCCAAAACGGAAGTGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATACCAATGCCTCAGCTGCCA
TTGCACCCAAGGATAATCTTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT
GAAGTCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGGACACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTGAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTCTATCCGGGAGCGGCGTCGCACCCTCCCCTCAGGGTATTGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTTATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAACCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCTTTCTCCGCA
GGGCCCAGGAAGTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGGCGTTGAT
GCTGCTGCACTTCCGGTTCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTTGCACTGACTTTCTGAC
CCATCCACCTGTTTTTTTGCAGATTGTCTATGAATAAACGGTGCTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRRLQCFQPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFTDVIRERRRRLPTQGIDDFKDKAKSKTLDLDFIDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCATTGACGTGG
TACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTTCGCGTGTGTCTTGC
ATAATTTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACATGGACAGAAGCTGGTGGACAGTTTT
GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGCAAGCAATGTGTTGCTT
GTGATTGCAACATTTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAATCTTCCTCATGTACCTGTTTCCTC
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGAGTGGAGCAGCCAGCAGGCGCCCAACATGCTCTGTCTGTGCCTG
TACGTGCCGGTTCATCGGGGAAGCCAGACCGAGTTCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC
ATTTTCAAGCTCAGTGTCTTCATCCCCCAGGAATTCCTCCACCTACCGCCAGTGGAGCAGAAAAATGTACAAGCT
GGAGATAAGGACCTTGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTTTAAGATTTTGGACAAAAAATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTCGGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATCTCAAGAGCATGGATAAAAAACGGCAGCATGACCATCGACTGG
AACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAAACATCCCCGAGATCATCTCTACTGGAAGCATCCAGC
ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGGTGGGGCAGGGGCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGG
TCACTCTGGCGGGGCAATGGCATCAACGTCTCAAAATGCCCCGAATCAGCCATCAAAATTCATGGCTATGAGCAG
ATCAAGCGCCTTGTGGTAGTGACAGGAGACTCTGAGGATTCACGAGAGGCTTGTGGCAGGCTCCTTGGCAGGGGCC
ATCGCCAGAGCAGCATCTACCAATGGAGTCTGAAGACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGTGGCCGCTTCTACAAAGGCTATGTCCCCAACATGGCTG
GGCATCATCCCTATGCCGGCATGCACCTTGACGTCTACGAGACGCTCAAGAATGCCTGGCTGCAGCACTATGCAGTG
AACAGCGGGGACCCGGCGCTGTTGTGCTCCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTAC
CCCCTGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC
TTCAAACATATCCTGCGGACCCGAGGGGGCCTTCGGGCTGTACAGGGGGCTGGCCCCCAACTTCATGAAGGTCTATCCCA
GCTGTGAGCATCAGTACGTGGTCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTCCGGGTGACGGGGGAGGGC
CGCCCGGAGTGGACTCGCTGATCCTGGGCGCAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACACT
AAGCTGTCTCGAGCCAAGCTGTGAAAACCCCTAGACGCACCCGAGGAGGGTGGGGAGAGCTGGCAGGCCAGGGCTT
GTCCTGCTGACCCAGCAGACCCTCCTGTTGGTCCAGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTGCAGCAGG
CTCCGGGCTCACATGTGTAAAGCAGGACATTTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCGGGCT
TAGTCTTCCATTTACCCTTGCAGCCAGCTGTTGGCCACGGCCCCCTGCCCTGGTCTGCCGTGCATCTCCTGTGC
CCTCTTGTGCTGCCTGCTGTCTGAGTAAGGTGGGAGGAGGGCTACAGCCACATCCACCCCTCGTCCAATCCC
ATAATCCATGATGAAGGTGAGGTACGTGGCCTCCAGGCCCTGACTTCCCAACCTACAGCATTGACGCCAATTGGC
TGTGAAGGAAGAGGAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATG
CTTGGGAGTGCAGGGGCTCGGGCTGCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTCTCTGAGCT
GGCCTGGACCTGTGAGGATGGGCCCCACCTCAGAACCAACTCACTGTCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTTGGGGCAGGGCAGAGCGTTTGTGTCTTGGGAGGGAAGGAAAAGGTGTTGGAGGCTTAATATATGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCTGTCCAACCCAGCAGGGGCGCAGC
GGGACAGCCCCACATTCCACTTGTGTCACTGCTTGGAACTATTTATTTTGTATTTTGAACAGAGTTATGTCTT
AACTATTTTATAGATTTGTTAATTAATAGCTTGTCAATTTCAAGTTCATTTTATTATATTTATGTTTATGTT
GATTGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGAAGGGGGCCTTGGGCCGCTGCAGTCACTCT
GTCCAGAGAAATCCTTTTGGGACTGGAGGCAGAAAGCGGCCAGAGGCAGCAGCCCTGGCTCCTTTCTTTGGCAG
GTTGGGGAAGGGCTTCCCCCAGCCTTAGGATTTAGGGTTGACTGGGGGCTGGAGAGAGGGGAGGAACCTCAAT
AACCTTGAAGGTGAATCCAGTTATTTCTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAG
TGAGGTGCTCTCACTGTGAATTTGTGGTGGGCGGGGCTGGAGGAGAGGGTGGGGGCTGGCTCCCTCCCTCCAGC
CTTCTGCTGCCCTTGCTTAACAATGCCGGCAACTGGCGACCTCACGGTTGCACTTCCATTCCACCAGAAATGACCTGA
TGAGGAATCTTCAATAGGATGCAAGATCAATGCAAAATTTGTTATATATGAACATATAACTGGAGTGGTCAAAAAG
CAAAATTAAGAAAATTTGGACGTTAGAAGTTGTCATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAGCTGAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAEQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQIVQAGDKDLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRCLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIAK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTGCGCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCTGGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCCTGTGAGGCTCCCCGATGGTTCCCC
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTTCTCGGAAGTCTCCAATAC
CAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
ACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGTAAACTCAAAGGCTTCTCTGTGTGT
CTCTTCTTTCTTTGCCATCAGCTGGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT
GTGCCTTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTAC
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTATGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCAC
AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRKLVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTSEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCTGGTGGTGGTTATGGGGGTCTGCCCTGGAGGGCCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCTCTGGAATCCAGGAGGACCATATG
GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCTTACGGTGCCAGCAGCCT
GGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCCAGTTACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTATGACCCCAACCATCTGTGGAGAGTGGAGTGCACCAGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTCTTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCTGCATCATAGCCACCA
AATAGTGAGGACCGGGGTGAGGCCACACAGATAGGGGCCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTCATCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTCTCAAAGTGGAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAAT
CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAATTGCATTTTTTTTT
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVSDSH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYPANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGCCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTCCCAAACCTCGGGACC
AACTTGCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTAACTCTGAACATCCGCAGCCCCG
TCTGGACCCCTAGGTCTAATGACTTGGCAAGGGTTCCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGGGA
AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGTCTGCGGCCCTCGTCCGGGCAGTGGCCCTT
TGCCTGGGGAGTCTTCTCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCCGTCTAATCACTGGGAGCCGGGGGAAAAATCCTTTCCCAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG
TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCAAGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAATCGGTATCCAGGAGCAGCTGGGGGAATATTAATCGGTATCCAGGAG
GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCTCCTGGAGTTCTCCGC
CCTCCTGGCTCTTCTTGAACATCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGGCTAGAGCACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCTGTCTCCGCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCACTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGSPNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPFAGGSAVQRWPPSWGLPAMDSWPPEDPWQMAAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRNSNLGAGGKILSORP
PWSLIHRVLPDHPWGTINPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSWGNNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVL RPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCCCTTGTCTCTCTTGACCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATTCACAGCCATTGAAGCCTGTGTCTTCTTGCCCC
GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCTC
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MGSGPLVLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC
CAGGTGCCCCGTGCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCGGGCTGGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
GCACTGTTGGTGCGBAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGAAGAAGGTACTTCAA
AGACTCTGCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSTSSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLLALIVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTGAGAGGTTTTGTTTTCTCTTAGTTCTGTGCTGCTGCACCAG
TCAAATACTTCTTCATTAGCTGAATAATAATGGCTTTGAAGATATTGTCATTGTTATAGATCCTAGTGCCAGAA
GATGAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA
TTTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAACATGAA
AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCAGGTAGAGATGAACCATACACCAAGCAGTTTACA
GAATGTGGAGAGAAAGGCGAATACATTCACTTCACCCCTGACCTTCTACTTGGAAAAAACAAAATGAATATGGACCA
CCAGGCAAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTC
TACCGTGCTAAGTCAAAAAAATCGAAGCAACAGGTGTTCCGCAGGTATCTCTGGTAGAAAATAGAGTTTATAAGTGT
CAAGGAGGCGAGCTGTCTTAGTAGAGCATGCAGAAATTGATTCTACAACAAAACGTATGGAAAAGATTGTCAATTCTTT
CCTGATAAAGTACAAACAGAAAAGCATCCATAATGTTTATGCAAAGTATTGATTCTGTGTTGAATTTTGAACGAA
AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAAT
TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTCTCATTGCTGAAGATCAGTCAA
AGAATTTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCTAAATCGAATGAATCAAGCAGCA
AAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTA
AATAAGCTAATCCAAATAAAAAAGCAGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
ACTTCCATCTGCTCTGGAATTAATATGCATTTGAGGTGATTGGAGAGCTACATTCCCAACTCGATGGATCCGAAGTA
CTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTCTTGATTGATGAAGTGAACAAAGTGGGGCCATTGTTTCAT
TTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT
TCAGATGAAGCTCAGAACAATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAG
TCCTCTCAGCTCGAAAGTAAGGATTAACTGATAGTAATGCCTGGATGAACGACACTGTCAATTTGATAGTACA
GTGGGAAAGGACACGTTCTTCTCATCAGATGGAACAGTCTGCCTCCAGTATTTCTCTCTGGGATCCAGTGGAACA
ATAATGGAATAATTCACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCCAGGAACTGCAAAGGTGGGCACT
TGGGCATACATCTTCAAGCCAAAGCGAACCAGAAAACATTAACATATTACAGTAACCTTCTCGAGCAGCAAAATCTTCT
GTGCTCCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCAGCCCAATGATTGTTTACGCAGAA
ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCATTGAATCACAGAATGGACATACAGAAGTT
TTGGAACTTTGGATAATGGTGACGCGCTGATTCTTCAAGAAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT
ACAGAAAATGGCAGATATAGCTTAAAGTTCTGGGCTCATGGAGGAGCAAACTGCCAGGCTAAAATTACGGCCTCCA
CTGAATAGAGCCGCTACATACCAGCTGGGTAGTGAACGGGGAATTTGAAGCAAAACCCGCAAGACCTGAAATTTGAT
GAGGATACTCAGACCACCTTGGAGGATTTGAGCCGAACAGCATCCGAGGTGCATTTGTGGTATCACAGTCCCAAGC
CTTCCCTTGCCTGACCAATACCCACCAAGTCAATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTT
ACATGGACAGCACCAGGAGATAATTTTATGTTGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT
GATCTAAGAGACGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCAACAAAGGAGGCCAACTCCAAGGAA
AGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCCACATATTATTGCCATTTAAAGTATAGATAAA
AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACTTTGTTTATCCCTCAAGCAAAATCCTGATGACATTGAT
CCTACACCTACTCCTACTCCTACTCCTACTCCTGATAAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTATTG
TCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTTC
AAGTAGACCTAGAAGAGAGTTTTAAAAAACAAACAATGTAAGTAAAGGATATTTCTGAATCTTAAATTCATCCCAT
GTGTGATCATAACTCATAAAAATAATTTTAAAGTGTGCGAAAAGGATACCTTTGATTAAATAAAACACTCATGGATA
TGTAATACTGTCAAGATTAATAATTAATAGTTTCATTATTGTTATTGTTATTGTAAGAAATAGTGATGAACAAAG
ATCCTTTTTCATCTGATACCTGGTGTATATTATTGATGCAACAGTTTCTGAAATGATATTCAAATTCATCAA
GAAATTAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAATAAACAACATTTGGAAAAAATAAAAA
AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIEDMVTASTYLFE
ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDfKNTIPMVTPPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRNLN
MNQAAKHFLLOTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGSEVLLLTGDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVDEAQNGGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNDGVYSRYFTAYTENGRYSLKVRHGGANTARLKLRLPPLNRAAYIPGWVNGEIEANPP
RPEIDEDTQTTLEDFSRASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN
FDVGKVQRYIIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIQVTLFIPQANPDDIDPTPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVI
VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCGGGCAGGGGTGA
CAACAGGTGTCATCTTTTTGATCTCGTGTGTGGCTGCCTTCTATTTCAAGGAAAGACGCCAAGGTAATTTTGACCCA
GAGGAGCAATGATGTAGCCACCTCTAACCTTCCCTTCTTGAACCCCCAGTTATGCCAGGATTACTAGAGAGTGTCA
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCTGTGGTTGGAGGAGAGAACCCTTGTGGGGCTGCGTTCTCTTAGCA
GTGCTCAGAAGTGACTTGCCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCTTCTGCTGTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAACTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCTTTAAA
AGTAGAGAAGCTGCTGTGTGGTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCCAGGGAAGCCCTTCCGTGGGGGCCCGGCTTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGTTTCCGCGGGGGCTGCTTGCCTGGATTTCCCGGGTGGTG
GTTTTGCTGGTCTCTCTGCTGTGCTATCTCTGCTGTACATGTTGGCCTGCACCCCAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGGCCAACAGCCCCACGGGGAAGGAGGGGTACCAGGCCCTCCTTCAGGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGG
AATGGGCGAGTACCAAGCCAGGATGCTGCTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC
CTGGCCTTCTGCACTCGCAGTGGAGTGGACAGGAGGATGAATGCTGGCTCAAGCTGGCCACAGATATGCAGCAGTG
CCTTTTCGATAGCTTTACTCTACAGAAGGTATACAGCTGGAGACTGGCCTTACCCGCCACCCCGAGGAGAAAGCTGTG
AGGAAGGACAAAGCGGATGAGTTGGTGGAAAGCCATTGAATCAGCCTTGGAGACCCTGAACAATCTGCAGAGAAGCAGC
CCCAATCACCGTCTTACAGGGCCTCTGATTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTAT
GAGCTCACCTTCAAAGGGGACCAAAACAGCAATTCAAACGGCTCATCTTATTTCGACCATTCAGCCCCCATCATGAAA
GTGAAAATGAAAAGCTCAACATGGCCAACAGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTC
CGGCGATTCTGCAGAAATTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG
AAAGAAGAAAATAATGAAGTCAAAGGAATACTTGAACACACTTCCAAAGCTGCCAATTTCAGGAACCTTACCTTCATC
CAGCTGAATGGAGAATTTCTCGGGGAAAGGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGTCCTTCTC
TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAG
GTATTTTATCCAGTTCTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGAA
CAGCAGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGATTGGGATGACGTGTGAGTATCGGTGAGAC
TTTCATCAATATAGGTGGTTTATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGAGCGCTGTGCGAGGACTCTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG
CTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG
GTGTTTCAGGCACGAGATAGAGGCTCACCTTCGCAACAGAAACAGAGACAAGTAGCAAAAAACATGCACTCCACAGA
GAAGGATTGTGGGAGACACTTTTTCTTCTTTTGAATTAATGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA
GGACGACAAAAGAAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTGTTGGGCTTTTACAACAGA
AATCAAAATCTCCGCTTTGCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTCTGACAAAGGCAGAAATGCTTGTG
AGATTATAAGCCTAATGGTGTGGAGGTTTTGATGGTGTTCACATACACTGAGACCTGTTGTTTTGTGTGCTCATTGA
AATATTCTGATTTAAGAGCAGTTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAATATCAGAAAGGCAGGAGAGAGATAGGCTTATTATGATACT
AGTGAGTACATTAAGTAAATAAAATGGACCAGAAAAGAAAAGAACCATAAATATCGTGTCAATTTTCCCCAAGAT
TAACCAAAAAATCTGCTTATCTTTTGGTTGCTTTTAACTGTCTCCGTTTTTTCTTTTATTTAAAAATGCACT
TTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCCTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTAC
ATTTTTATATTTTTTAAGAAGATACTTTGAGATGCATTATGAGAACTTTTCAGTTCAAAGCATCAAATTGATGCCATAT
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCATTGAGACATAGGGAAGGAATGGTTTGTACT
AATACAGACGTACAGATACTTTCTCTGAAGAGTATTTTGAAGAGGAGCACTGAACACTGGAGGAAAAGAAAATGAC
ACTTTCTGCTTTACAGAAAAGGAACTCATTGAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAACCATTTT
CTCCTCAGAAGTAGGACCGCTTTCTTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACATCTCTTTTC
AAAACAGGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATTTGT
GAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTTTGTACATGTTATCCACCCAGGCCAGGTGGAAG
TAAGTGAATTTTTTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATTGCATTTTATTACCATTT
CAAATTTTTTAAAAATAAATACAGTTAATACATAGAGTGGTTTCTTCAATCATGTGAAAATTATTAGCCAGCACCAG
ATGCATGAGCTAATATCTCTTGGTCTTCTGCTTCTGTTGCTCAGTAAACTCATTGTTTAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTGTACTGTTAGTTTATGAAATTAATTAACACAGG
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTTTGTGGATATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNEKLNMAN
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFSQY
NPGIYGHHDVAPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDDIKGWGGEDVHLYR
KYLHSNLIVVRTPVRLGFLHWHEKRCMDLTPEQYKMCQMOSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLNRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCTGTGTC
TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTGTCTGACAGCTCC
AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCCTTCGTCCC
AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
GATTTTGATTTATGGAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAATCGTGATTA
ACTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGACTCTGAAG
AAAACACGGAAGGTACTTCTCTACCCAGCAAGAGTCCCTCAGCAGAACATACCCCCGATAAA
ACAGTCATTGAATATGAATATGATGTGACAACCACTGACATTTGTGCGGGCCTGAAGAGCAGGA
GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
TCTTGGGCCCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCTGACTGGGA
TCCCCAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG
AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTATA
TGTGCAGATGGAAACTGATGCCAACACTTCTTTTGCCTTTTGTTTCCTGTGCAACAAGTGAG
TCACCCCTTTGATCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTTCCAGTTTGTGAGTGT
CTGTGAGAATTACTTATTTCTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTTCATGCATGTA
GGTCTCTTAACAATGATGGTGGGCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSISSQIGPPEVALTTDEKSSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRIHVGKEKHPANLILYGNEDKRFFVPAEK
IVINFITLNISSDKISHQDMSLLGKSSDVSSLNPPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPDPKTVIEYDYVRTTDICAGPEEQELSLQEEVSTQGTLLSQA
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
CTGGGAAGATGCGCGCCCGTGGACCTTCACCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
CAAGCCACCCTCAGTCCCACTGCAGTTCTCATCTCGGCCCAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCTGAAGCACATC
ATCTGGCTGAAGGTCATCACAGCTAACATCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
CCCACCCGCTGGTCTCCTCAGTACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTCTGTA
TAAGCTCTCCTTCCTGGTGAACGCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
ATTGATGGTCTGTTGGACTCTGTGCTTCCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGT
TCCCTCCAGTGAAGCCCTCGGCCCTTGTTCACCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTATACTCAACTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
CCTTGGGATTGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGATGGCAGCCATCAGGGAAGGCTGG
GTCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPKIGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMTLDNIPFSLIVSQDVVKAAVAVALSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTkdalvltPASLWKpssPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACCTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCCTTCTGGGCCTGCCCCGTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTCATCCTTGGAGGCCTCCTGGGATT
ATTCTGTTCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCCTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTTCTGCTCATCCCAGAGAAATCGCTCCAATACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCCTACAGCCTGACAGGGTATGTGTGAAAGAACCAGGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
TCACCTTGCTGCTCCCCCTGCCCTAAGTCCCCAACCTCAACTTGAAACCCCATTCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAACCCACTAATCACA
TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCCTCTTGTTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYYLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSSAISLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAVNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNYSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTC
CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCCT
GCCCACGCTGCTTCTTGGCCCTTCTCCGACCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCTCCGTGTCTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGG
ACCAGCGGCCTGACCCTGGGGAAAGCATGTTCCCGAGGTGAGGGTCCTCTCCTCCTTGCTGGGA
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCCCACCAAAGTCTGCCAGCACAAACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCTGCCCAACAGTGTGTCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCGGCCTCACAACTGCCCCGAACAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCTGCAAAGATGAGGCAAGTGAAGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGGTGAGACATCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG
AGAGGGCCCGGCGACCCAGCCCCACTGGCCTCAGCGCCCTCTGAGCTTCATCCCTCGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAAGTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
CCCTTGCCCTGCATCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCCAGAGG
ACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGGTC
CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCACGAAGGTCACT
GGAACGTCTTCCTAGCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA
TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTIYPCRHPKAVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCAGCAGAGCCTCTCC
 GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCATCCGTCACCTCTCCTGTCA
 TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTTCTGAGTC
 TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAG
 GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCAGTTCTTCAGGGG
 CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
 AAGGCAGGACAAAATGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAATACTACT
 GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
 ACAGGTGTGAGCACTGGGCTCAGTTCTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
 TCTGTCACTCCTCGGCTGGTTCCTCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTTGTCC
 ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA
 CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
 GAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA
 TTTTTTGGCATTGTTGGACTGAAGATTTTCTTCTCCTCAAATTCAGTGGAAAATCCAGGCGGAACCTGGACTG
 GAGAAGAAAGCACGGACAGGCAGAAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
 AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGGTG
 CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAACATTA
 CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCGTGGGAGTGTCCGGGATGATGTGGACAGGAGGA
 AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTT
 ACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAATAGGGGTCTTCTCTGGACTA
 TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCCGTTTG
 AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCCTATAATGAGCAAAATGGAACCTCCCATAGTCACTGC
 CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGCCAAAGGGCCTCTGCAATCCCAGAGACAAGCAACAG
 TGAGTCTCTCTCACAGGCAACCACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
 TCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCA
 GATGAAGGGGACTGGCCTGTCCACATGGGAGTCAAGTGTCTATGGCTGCCCTGAGCTGGGAGGGAAGAAGG
 CTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG
 AACCGTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAAGTAGACAAGGAATGTGAATAATGCTTAG
 ATCTTATTGATGACAGAGTGATCCTAATGGTTTGTTCATTATATTACACTTTCAGTAAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSGQWQVFGDPKPVQALVGEDAAFSCLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPQYQGR TKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPRETA KWKGPGQDLSTDSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKT VTHRKA PQEVP HSEKRF
TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHL YFT
LNPRFISVFPRTPTPKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAI PETSNSSESSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
TACCCTGGCCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGATCGGGGAGA
TACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCCCCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTGAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAAGTTGATGCAAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAACGCTGTGAGGGTTT
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAG
CTTCTGCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAATAAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRENILIPGTLESGCPQNLTCVWPWACEQGTTPPMISWIGTSVSPDPSTTRS
SVLTLIPQPQDHGTSLTQVTFPGASVTTNKTVHLNVSYPPQNLMTVFQGDGTVSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLSQKATSGVTQGVVGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTE
DANAVRGSASQGPLEPWAEDSPDQPPASARSSVGEGLQYASLSFQMKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGAGGGGAGGGAGTGAAGGAGCTCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTTGA
TGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTACCCCTAT
GGCCAGCGGGAATTCACTGCGGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
GGAACATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIPIVVYDFGDAQKTASYISPYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTACAGCTGGTGGCCGCGCTCT
CTGTACCCCTGGGCCTCTTTGCAGTGGAGCTGGCCGTTTCTCTCAGGAGTCTCCATGTTCAAC
AGCACCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTTGTCTTCTGCAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTTATTTCGTACCGTCTTTGGGCTGAAAAAGAAACCCTTCTTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCCCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

APP ID=10063569

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTTCGTGTTCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCTTCCCGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTACCTAAGTCCCCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACGCCCCGAACCACAGCCCCC
CACCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCGCCGGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACACTAC
CATAGTGATACCCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGGCTC
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCCTGTCTGTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCAACCATTGCTGATG
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
CGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGAGACGGGGTCTCCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGCGGGCCGAGAGCATGTGCTGGATCTGTTC
TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCC
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT
CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA
GTGGAGTATTTTCCTTTGTGAATACCCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTTCAT
GGAGATTCTCCTCAATTCTTCCTTTATTTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGTGAGTGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTGTCAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTTATAATCCTGGTGTCTATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTCTTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT
TTTTCTGAAAATCCCTTTTTCTGGTGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGPPTVGIHGDSPOFFLYSAFMTLVIILLHVFVGIVFFDGCCKKKGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAAGTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAAAC
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTTGTTCCTGGGAGGCATCTGCCACCAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCAGCAGGCCCGCTCCCAACTCCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAATGGAATTCA
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTGGTGATACATGTGAATCTTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT
TACCTGAAAATATTCTTGAAATTTTCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGGAAAAGTGGAAA
AAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTLISSEE
LPQIFTSLLIHSFLPGGILPTSQAGANPDVQDGSLLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTGCCTTGCCCCACCCAGCCAGCCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTGACCATG
 GTCCCTGCCTGGCTGTGGCTGCTTGTGTCTCCGTCCCCCAGGCTCTCCCAAGGCCAGCCTGC
 AGAGCTGTCTGTGGAAGTTCCAGAAAATATGGTGGAAATTTCCCTTTATACCTGACCAAGTTGC
 CGTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTGAGGGGACTCAGGCAAGGCAACTGAG
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCTGCTGGTGACCAGGGCCCTGGACCCGAGAGGA
 GCAGGCAGAGTACCAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGTCTTGTGGGTCCAC
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC
 AGAGCTCGGCTGAGCCGGGTACCAGGCCTGGCATCCCCCTTCTCTTCTTGGGCTTACAGCCG
 GGATGAGCCAGGCACAGCCAACCTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCAGC
 CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
 CCAGGCCCTCAGGCCACAGGCCACTGCCACCGTGAAGTCTCCATCATAGAGAGCACCTGGGTGT
 CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTATACCCGCACCACATGGCCCCAGGTA
 CACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCGGGACCCCTTGAAGTGAA
 TGCAGAGGGAAACCTTACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCC
 AGGTGCGGGCTCAGAAATCCCATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTG
 ATGGATGAGAATGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
 CAGTCCACCAGGTACTGAAGTACTAGACTGTGACAGAGGATGCAGATGCCCCGGCTCCCCCA
 ATTCCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC
 CAGGTGGACCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT
 CCTGCTTCTGGTGTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGTGAAG
 TCGAAGTCGCAGTCACAGATATCAATGATCACGCCCCTGAGTTCATCACTTCCCAGATTGGGCCT
 ATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
 CCTCGAGCCCGCTTCCGCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGACTTTTG
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCCC
 AGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGATGCCACCCCCAAGTTGG
 ACCAGGAGAGCTACGAGGCCAGTGTCCCATCAGTGCCCCAGCCGGCTCTTTCTGCTGACCATC
 CAGCCCTCCGACCCCATCAGCCGAACCCCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
 CTGCATTGAGAAATTCCTCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG
 ACACCTACACGGTGCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
 TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGACCCAGCAAGGACCCCGATCT
 GGCCAGTGGGCACGGTCCCTACAGCTTACCCTTGGTCCCAACCCACGGTGCAACGGGATTGGC
 GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCAGT
 GAACACATAATCCCCGTGGTGGTGCACCAATGCCAGATGTGGCAGCTCCTGGTTTCGAGTGAT
 CGTGTGTCGCTGCAACGTGGAGGGGAGTGATGCGCAAGGTGGGCCGATGAAGGGCATGCCCA
 CGAAGCTGTGCGCAGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAATCTTCTCATCCTC
 ATTTTACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGGCCCT
 GAAGGCGACTGTCTGAATGGGCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
 TCCCCTGGGAGAGAGCCAGCACCCAAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCCCTCA
 TCTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCAACTT
 TATGGACTGCCCATGGGAGTGCTCCAAATGTGAGGTGTTTGCCCAATAATAAAGCCCCAGAGAA
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDGQASGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDYAAPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPLRAGONILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADADLEPAFRLMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWWEP
REHIIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTTLVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCCACGCTTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
TCTCTTACCTTCAAGTCCCCTTTCTCAAGATCCTCTGTCTTTGCCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
GCATAGTCACCAACTCTGAGTTCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCCACCACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCTCCAGTGGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCAACCAATTCTGAGTCCAGCACACCCTCC
AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGTCCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCAACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCACACAACCTTCCCATAGTGATCTACTGCACTGAGTGAGGCAAAGCCTGGTGGGTCCCTGGT
GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC
TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCTAACTGGTTCTGGAGGAGACAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
ACAGCGGGCCCCTGAGCAGCCCCGGAAGCAAGTGGCGCATTTCTCAGGAAGGAAGAGACCTGGGCA
CCCAAGACCTGGTTTCTTATTCTATCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCTGAA
AATCTTGAAGAAGGTATTCTCACCTTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATAT
TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACAGACAAAGAGAAGCTGTGCTTG
CCCCGGGGTGGGTATCTAGCTCTGAGATGAACCTAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNLSLRNTFNTAVYHPGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVSII
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCTCCGCGTTACGGGATGAATTAACGGCGGGTTCGCGACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACAGGTG
GGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCAATTCCAGGGGAAGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTCATGGAAGGGATTTACTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGAAGCATTT
TCATGCTGAGTCCCTTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAGAGT
GATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCTCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGTTGGTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCATAGGAAATGGAAGGATGACAAGAGCCATTTGAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCAACTCCTCATATCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAACATTCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTTACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAACCGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCATTCCACCTTGCAAGTCTGAACCTCAGGGTC
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTTCAGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTCATGACTATGTGCAATATTTCTTACT
GCCATCATTATTTGTTAAAGATATTTTGCACCTAATTTTGTGGGAAAAATATTGCTACAATTTT
TTAATCTCTGAATGTAATTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGCCGAGAATATTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGFIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCIMRYSYLRLLEKICLKASLKGVPFGFWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFFEGTDLTENSRSRNSAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFFPREIHVHRYPIDTLPTSKEQLWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVLVVKLLSILYWTLSFAMCLLIYLSLVKWFITITIVIFVLQERIFGG
LEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCTATCTGTTTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCCTTTTGTGCTCATTGTTGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGTTTCTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA
GTTTGTGAAGGACTCCATTGCAGGGGGCGTGTCTCTTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTTCAGTTCCAGATTTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTTCTCTCATTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCTCAGGCTGGTTCCCCAGCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGCGCGCTGGCTTCTA
TTTTACTCGGGTTACTCTGTGGTGCCTGTGTGGTGTGTCATGGGGATGATAATTGTTTTCTTC
AAATCCAAAGGGAAAATCCAGGCGGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG
TTTCTGATCTGAAAATGTAACCATAGAAAAGCTCCCCAGGAGGTGCCCTCACTCTGAGAAGAGA
TTTACAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACA
ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTC
ACATTCAATCCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCTCT
GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCTGCT
TGACATGTCAGTTTGAAGGCTTGTGTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATATTCTATATGTCCAGTGTCTGGGATGAGACAGAGAAGACCCTGCTTAAAGGGC
CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCT
CCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC
TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA
CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGG
ACAGTGATTCCTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
TGAGGGCACAGTGTGTTGCTAATGATGTGTTTTTATATTATACATTTTCCCACCATAAACTCTGTT
TGCTTATTCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC
TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC
CTCATTAACACAGACACAAAATTTCTAAATAAAATTTTAAACAAATTAACATAAACAATATATTTA
AAGATGATATATACTACTCAGTGTGTTTGTCCCAAAATGCAGAGTTGGTTTAAATTTAAAT
ATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGTGGGAACCCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCATTTACAAC
TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTGAAAAATGCATTTTATAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
ACCGAAGAAGTAAACTCTAGGTGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCACTCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
TTGTGAGTGTGCTCACTGTTTACAACATATAAGAACCCGCCAGATGGACTGCTTCTTTGGAGTAACA
ATAAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACCCCATCACA
TGACTATGATATTTCTCTTGCAAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTC
TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTGTGACAGGATTTGGAGCACTGAAAAAT
GATGGTTACAGTCAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACTTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
GCCAGGGTGACTCTGGAGGACCCTGGTTAGTTTCAAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTGTTATACTAGAGTTACGGCCTTGCGGGACTG
GATTACTTCAAAAACCTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTGT
GGTGTGGAGGCCATTTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCA
ATAAACTGTTTGCTTGATGCATGTATTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA
GATCAACTCTGTCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTGTGAGAATTTTGACTTGTGACATAAATTTGTAAT
GCATATATACAATTTGAAGCACTCCTTTCTTCAGTTCTCAGCTCCTCTCATTTTCAAGCAATATCCATTT
TCAAGGTGCAGAACAAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCTACATTTTATTGGCACAGAA
AAGTATTAGGTGTTTTCTTAGTGGAATATTAGAAATGATCATATTCAATTATGAAAGGTCAAGCAAAGACA
GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAAGTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
ATATATCCTTATTTTCATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCT
ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTTATTTAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELSSPVPYTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSSGGLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCCTG
CCCCGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGGCCCAGC
GCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
ACGGGGGTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA
TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTCCT
GTGGTTTCATCCGGTCATGATTGCTGTTTGCTGTTTCTTATCATTGTGGGGATGTTAGGATATTG
TGAACCGGTGAAAAGAAATCTGTTGCTTCTTGCAATGTTTGGAAAGTTTGGCTTGTCAATTTCT
GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTATGGTTCCAGTACAATGGTCA
GATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
TGCTTGGAAATTTTTTTCAGAGAGAGTTTAAAGTGTGTGGAGTAGTATATTTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCCAGATTCTGCTGTGTGTAGAGAATCCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTGTGGGAAGAAAATGTATTCCTT
TTTGAGAGGAACCAACAACCTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
ACAGACCAATGATGTCCTTGAAGAATGACAACCTCTCAGCACCTGTCTGTCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
TTGAGATGGAGGAGTTATAAAAAGAAATGTCACAGAAGAAAACCAAACTTGTTTTATTGGACT
TGTGAATTTTGTAGTACATACTATGTGTTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA
TAACACCTAAGCATATACTATTCTATGCTTTTAAATGAGGATGGAAAAGTTTCATGTCTAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCGCA
TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
CTCAGCGATCTATTCTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACT
ACCTAAATGTGATTTTTGCTGGTTACTAAAATATTCTTACCACTTAAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCTGTATAGCATCATT
ATTTTTAGCCTTTCTGTAAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
TGTTATTTAAATACTTAACCACTAATTTTGAAAATTACCAAGTGTGATACATAGGAATCATTATTC
AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTGAGA
AAGACTTGTATGCTGTTTTCTCCCAATGAAGACTCTTTTGGACACTAAACACTTTTTAAAAA
GCTTATCTTTGCCTTCTCCAACAAGAAGCAATAGTCTCCAAGTCAATATAAATCTACAGAAAA
TAGTGTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAAAACATGTGACAATTTAGAGATT
CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTTTTACAA
GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGATTTTGTGTTAT
TTCTCAGAAATATGGAAGAAAATTTAAATGTGTCAATAAATATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCTCATCTCTTCTCTGATAAAGCCCCCTACCAGTGCT
GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTCACCCGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAAGTTGGATGAGTCTCAAGGTCTCTCTATGTGACAACCATGTGAATGGGGAG
TGGTACCACCTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
TCAATGGGAACGTCTGTCTCTGGAACACCACGGTGAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGCTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAACGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCTACCCTGT
GAGTGTGGGGTTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTTGCCCTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCTGGTGTGT
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGTTGGTGGCTGGAGCTCTTCTGACCAACACCTCTGCGGAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTTCTCTCTCAAGACATGTGGTACAGTGGTCTGATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCCGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
ACCTGCGAGTTTCCACGCCTGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGGATCTTCCATTCACTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
GGGAAGCTCTGCCACCCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTGCTTTGCCACCCCACTCCAAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG
TCCTGTCTTCAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTTGTCTGTGGAGTG
TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGCGAATGCGTCTGGGGCAGGAGGAGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGGCGGCCGATCCGCATCGACTGGGAGGACTAGTTCTGAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACCGCCCTTAAGAACATCTGCCAACAGC
TGGGTTTCAGACTTCACACTGTGAGTTCAGACTCCAGCACCAACTCACTCTGATTCTGGTCCATTAGTGGGCA
CAGGTCACAGCACTGCTGAACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAATAAACTGTCCA
CCCAGAAAGACACTCACCCCATTTCCCTCATTTCTTCTACACTTAAATACCTCGTGTATGGTGAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTTCAATTCAAATGCAGACTAA
TTATAGGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
YCGHFDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDEVEGCHNNNGGCSHSLGSEKGYQCECPRGLVLSEDNHTCQVPVLCKSNAIEVNI PRELVGG
LELFLTNTSCRGVSNNGTHVNILFSLKTCGTVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEOPYREALPTLKLRLDSLYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYILIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVKGDKHEVFLHCRV
LVCGVLDERSRCAQGGCHRRMRGAGGEDSAGLQGGTLTGGPPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCCACTGCAGGTGCTGGACTCGGCCAC
AGGGAACCTGGTTCTCTGCCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAAACAGCCAG
GAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCCTG
TGGGAAGAGCCTGAAGACCCCCGTGTGGTGGGTGGGGAGGAGGCCCTCTGTGGATTCTTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCAGCTGTGTGGAGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
GCCCCTGCTTCAGGAAACATAACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAG
CTTCCCATCCCTGGCTGTGGCCAAGATCATCATATTGAATTCACCCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTTCCTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA
GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGT
ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
CTCGGGGGGGCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCCCTTTCAGTGCTGGGAGCCGCTTCCTTCCTGCCCTGCCACCT
GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTTGGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCCTCGCAGCCCAGAGGCGCCAGAGGAAGTCA
GCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAGCC
CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTCACCCATCCCCAA
GCCTACTAGAGCAAGAAACAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCAATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
AAAA

FIGURE 112

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVVLIKVILDKYYFLCG
QPLHFI PRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAE TACRQMGYSRAVEIGPDQDL DVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGEEASVDSWPWQVSIQYDKQHVC GGSILDPHWVLTAAHCFRKHTDVFNWKV RAGSDKL
GSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDT CQGDSGGPLMYQS
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWK AEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAACCTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAATC
TGTTTTTTGTTCTCTTGTAACTAGCCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCTCGAAG
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGGCTTTCAAGGTGGGC
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTTCATAGGCGATGGCTCCCACTGCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCCGCCCATCTAACCTTTTCATGTCTGCACATCACCTG
ATCCATGGGCTAATCTGAACCTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCATTCTCCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTGTGCCAAATTATGGGTGAGAAAAGATG
GAGGTGTTGGGTTATCACAAGGCATCGAGTCTCCTGCATTAGTGAGCATGTGGGGGAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAATCCA
AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCGAGGCGGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA
TGTTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:

amino acids 1-15

10063569

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAATTGATCCTGTGACCAGAAGTGAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAAACTCAGATTAAAGTGATTCCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTCTTAAAAATTCCAAAATCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCTGCCAACGAAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCCT
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTILIVLFWGSKHFWPEVPPKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCGAGCTTCTCGCAGGCGGCA
GGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCTCCTGTCCATCCTGGGGCT
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCC
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCT
GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTCAATGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTGCCAACATGCTGGTGACTAATTCTGGATGTCCACAGCTAACATGTACACCGCATGGGTGG
GATGGTGCAGACTGTTTACAGACAGGTACACATTTGGTGCGGCTCTGTTCTGTTGGGCTGGGTGCGCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAGAAGAAACCACTACAAA
GCCGTTTCTTATCATGCCTCAGGCCACAGTGTGCTTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCCGACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCACGACTATGTTAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
CCCAAAAAACAAGGAGATCCCATCTAGATTCTTCTTCTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTTTATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTCTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
TTCTCTAGTCAATAAACCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACCTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT
ACACTGTGATCTTAAAAGTTACCAACCAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTG
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCTCTTTCTGT
CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTAAGTCCTAAATATAGTTAA
AATAAATAATGTTTTAGTAAATGATACACTATCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTC
ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCA
TGGTGGCATAACCTGTAGTCCCAGCATTCGGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACCTGCACCTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCGGGAGTCC
AGCTGGCTAAACTCATCCCAGAGGATAATGCGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG
GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTTTAAAACTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGCTCATCCCTGTGAGCTGGGTGGCCAATGCCATCATCAGAGATTCTTATA
ACTCAATAGTGAATGTTGCCCAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTTGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTTCCCATCGCACAAACCAAAAAAGTTATCACACCGGAAAGAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG
ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAACTTTGATTTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGCAGAATGAGATA
TTAAACCCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTTATCATTTACTTCAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
ACGACATAGCATTATGTACATAGATGAGTGAACATTATATCTCACATAGAGACATGCTTATATGGT
TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCA
TTTATAATGAAGATTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAATCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT
CTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTTGTCAAGGGGCTTTCATTCAAACCTGCTT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG
AAAAATTTTTGTTTTGTATTTGAAGAAGATGATGCATTTTGACAAGAAATCATATATGTATGGAT
ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAGCAGAAAAATA
TGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAGTTGTCCTTTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTTTTGTTCTGTGAAAAATAAATTTCTTCTGTA
CCATTTCTGTTTAGTTTTACTAAATCTGTAAATACTGTATTTTTCTGTTTATCCAAATTTGATGAA
ACTGACAATCCAATTTGAAAGTTTGTGTCGACGCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTT
TATACATTTATATTAATAAATTGTACATTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFWGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCNEKSSSY
RYSIPSHRTTQKS YHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG
CCAGACGCTGACCACGTTCTCTCCTCGGTCTCTCCGCCTCCAGCTCCGCGCTGCCCCGCGAGCC
GGGAGCCATGCGACCCAGGGCCCCGCGCCTCCCCGCGAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCGGGATG
GATTCAAAGGAGAAAAGGGGGAATGTCTGAGGAAAGCTTTGAGGAGTCCTGGACACCCAACCTAC
AAGCAGTGTTTCATGGAGTTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATAAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTCACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTA
AAAAATATAAAGCTACCAATCTTTGTACAATTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAQVPGR
DGSPGANVIPGTPGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
 CTGGGTGCTCATCACGGGAACGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
 ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTA
 ACCGCCCCCTCCCCACCCCCAAAAAACTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
 TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGAG
 TGTTCGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT
 TCTTCCTGGGTGTGAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG
 GCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAGGATGGGTTTCAATG
 TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
 CTTTCTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
 ATCTCAGAAATTACAGGAGATACCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
 ATAACAGCCTTCAAAAACCTTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC
 CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTAATGGAATACGCAGACTCAAAGA
 GCTGATTCCTTAGTTCCAATAGAATCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAAT
 TACGGAACCTTGGATCTGTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG
 CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
 AGACTGCCGCAACCTGGAACCTTTTGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
 TCTTTGCTGGCATGATCAGACTCAAAGAACTTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC
 CTGGCCCTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGAATAAAATCAGTGT
 CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA
 TCGAAGCTTTTCAGTGGACCCAGTGTTCCTCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGAT
 TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG
 TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAAGTGGCTGAAAAGTT
 TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
 ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAGTACTACAGAGAGGTTTGATCTGGCCAG
 GGCTCTCCCAAAGCCGACGTTTAAAGCCAAAGCTCCCCAGGCCGAAGCATGAGAGCAAAACCCCTT
 TGCCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
 TTCCATAAAATCATCGCGGCAGCGTGGCGCTTTTCCTGTCCGTGCTCGTCATCCTGCTGGTTAT
 CTACGTGTCATGGAAGCGGTACCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
 GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTTATGTA
 GATTATAAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTA
 TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGCAACCATTTGTGATAAAAAGAGCTCTTAAAGCT
 GGGAAATAAGTGGTGTCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCCTCCCC
 TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCTTGTCCGTTTTAGTGCATTCAATACT
 GGTCAATTTCTCTCATAATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
 GAACTCCGGTTTAAATAATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTGCAATTT
 GTTTTAAGATAAAACTTCTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCGKMVYCESQKLQEIPSSISAGCLG
LSLRVNSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRITPVRI FQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLHLEHNQFSKLNALFPRLVSLQNLVYLQWNKISVIGQTMSWTWSSLQRLDL
SGNEIEAFSGPSVFCVQPNLQRLNLDNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN
WLKSEFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKI IAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKKRQSLKQMPSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTTCGCTACTGCTGAATGTCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT
AGCAACCTTTCCCTGGATCTCACA AAACTCGACTCCAATGCAAGGAGAAGCAGCTCTTGCTC
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTTATTGGCCAGTTT
TTAGCCAATCCAAC TGACCTAGTGAAGGTT CAGATGCAAATGGAAGGAAAAAGGAAACTGGAAGG
AAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
ACCAC TTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG
ACTGACTGCTTGATT CAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACC
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTPVKHYLVNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGAGAGCGGGCGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCGCTCGGGGGGCATCGGCGCGGCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGACTTTGATCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGTGGCTTGGCCCGGCGCTGACACCCTGCTCTCAGGCAGCACCAAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGCACATCATTAAATCAATAGCATGTCTGGCCACCGAGTGTTA
CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTTCGGGAGGCGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCTCAGCACCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
TCCTTCCTCCCCACCCTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT
CACGGGATACCATTCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC
TAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCTGGGCACCTGGCCTTTGTCTGCTCTCAGTG
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCCTTACCTTATATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCTCCTCTGCCTGCCCCACTGCACCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
AGCCCAGTCTTGGCTTCTTGTCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
CAGAACCAGGGCCTGGCCAGTGGATTTTCATGGTGATCATTAAAAAGAAAAATCGCAACCAA
AAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAARALVQQGLKVVGCAITVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSQVDICINNAGLARPDITLLSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHIININSMGHRVLPVSVTHFYATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGCTGGTGTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTTCAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATCCAGTTCTGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAACTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTTCATTTTCAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGAACGAGAAAATTAATTTCTCATGTATTTTCTCATTTATTTA
TTAATTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCACTAATTTTTGTATTTTGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTGCCTCGGCCTCCCAAAGTGTTATGATTACA
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTTCTTTGTGTTGGGAACTTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCACATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCTTG
TTCCATCCATGTTGCTGCAAATGACAGGATTTTCGTTCTTAATTTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGHGTLGESPMFPKRVCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCCTCAGAT
TTGAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCACAACAGTTTAGATGAATTTCCCTACCAACCT
CCCAAAGTATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
AAATTCCTTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTTTCCCTGTCCCGTAATCACCTTAGCACAAATCCCTGGGGTTT
GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTCAAG
GTCTCACTAGTCTAAACGCCTGGTTCTAGATGGAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
TTCCTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATTCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCCAATGCTTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGTATCTTT
GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
AGGTTTCGTGGGATGGCTATTAAAGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCCAAGGACAGTGGCCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTATATCTCTTGAAACTTGCTCTACCTATG
ACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATAGCCCGGCATTGGATCTATAACAGAAACAATTGT
AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTC
CCATGGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCCTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC
TTTGGCTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATGGCCCTTCTTGCTTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTCTTCTCAAGGAAGTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAAACTTCTTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAACACACCATATTTCTCCTAATGGAATGAATCTGT
ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATTGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNL PKYVKELHLQENNIRTITYDSLSKIPYL
EELHLLDDNSVSAVSIIEGAFRDSNYLRLLF LSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLRNSLT AAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDL DNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVMVPMETSNLYLFDETPVCIETETAPLRMYNPTTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIAL LALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNL YKNNHSESSNRSYRDSGIPSDSHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCAGGCTTCTTGGA
GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTTCCTAGAACAGAGGCTGCCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCGGGAGAAGTGGGCCAGGAGCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCTCCAGCCCGG
GTTTTGGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCCGGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
AGGCTACTGCCTGTCCCACCAACTGCTCTTCTTCTCTGGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCTACCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG
TGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCCTCCTGGGTTCAGC
AATTCTCTTGCCTCATCCTCCCAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT
TTTTATATTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCCAAAGTGTGGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
ATGGGGGCTCTCTCCCTAGATGGCTGCTCCTCCCAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCCTGGCAGAATAACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLARAQGAERTVSSEQRPMASLGLLLLLLTALPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLSHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQEGCGF
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCAGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTTCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTTCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTTATAGAAAATGCCATTAATAAATTAT
ATGAACACTATAACATTATGTATATTAATTAAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAAAA

FIGURE 136

MAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGSGVGIGDRFKIEGRAVVPGVKPQDWISAA
RVLVDGEEHVGFLKTDGSFVVHDIPSGSYVVEVVSPAYRFDPVVDITSKGKMRARYVNYIKTSE
VVRPLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSSGSSKTKGSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATCGCGTCAATTCCCCAAAACAA
GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAGGCTTGGCAGT
TTTTCTTACTCCTGTGGTCTCCAGATTTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCCTCAGC
CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTGCTAAGACTCTATCTGGACAGGGTATTTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAGGAGACAGAAATAGGAGGAAAGTGATGCTG
CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
CTTCCTTGCAATGATTGCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTA
ATGTATTTATTTTTTACTTGGACATGAACTTTAAAAAATTACAGATTATATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAATAACCTTGTAATTTCTAGAAGAGTGG
CTAGGGGGGTATTCAATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG
CCAGGAATCCTACACGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAA
AAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPG LTCWALTAEPGWGQNGATT CATNSHSDSEL RPEIF
SSREAWQFFLLWSPDFRPKM KASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLT IKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCCGCATCCAGCC
TAGCGTGTCCACGATGCGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCCTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCCCGAAACCCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCCTTGAGAGATGATTTTGTGTTTG
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTACCTTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCCCTGGCTTTGTCGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAGAAAGATAGTCTTTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAAGCATTTTGTGGAATATGATGGAACAACCTCATTTTTCGTGTCAGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTTGGATAAAGTATTAAGAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGGCTGGACCACATTGGCCACATTTAGGGCCCCAACAGCCCCCTGATTGGGCAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTGAAGGAGAGAGA
GACGCCCTTTACCAATTTGCTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
GGGCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTTAATCAGTTCTGCGTTTGAAAGGAAA
CCCGGTGATATCCGACATCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCCTATTCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG
AATGTCCGTCATATGAAAAGATCCTGGGTTTGAGCAGTTTAAATGTGAGAAAGATTGCATGG
GAACTGGATCAGACTGTACTTGGAGGAAAAGCATTGAGAAGTCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG
TTCTCACCTGCTCCTGCTCAGCGTCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA
CTGTCATCTCCTGGGTTTTCTCTGCTCTTTTATTGTTGATCCTGGTTCTTTTCGGCCGTTACAGT
CATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCTCTCGTGGCTGGCGGCAGGCT
GCCTTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCCTGGAC
AGGGGGCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCTCTGGGTGTCCCGACACAGGTG
TTCACATCTGTGCTGTCAGGTCAGATGCCTCAGTTCTTGAAAGCTAGGTTCTGCGACTGTTAC
CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCTGCCACCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTTCTTTCCAAGCGGACGTTTCTGTTGGAATCTTAGTCCTTGGCC
TCGGACACCTTCATTTCGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC
ACCCCAACCCTGCACAGCCCTCATCCCTCTTGCTTGAGCCGTCAGAGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTTGTATCAGGGCACAGGCTTCTCGAGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGCCCCATCTGGGCTCATGCTCTCTCTCCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAAGAATAAAGCGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPHFVEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDHHISGPNPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACAACCTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTAGCATTTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAAACTACATACGCCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCCCATTGAACGCCTTCCTCGCTA
ATTTGAACATAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLDSGNLIAVPDKNYIRPEI
FFALASSLSSASAEKGSPIILGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA
TCCAGGATCCTGTCCTTCCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGGG
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCTTGGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAAGTAGGTGTGGGAAATTTGAAGACGACATTGACAAGTCCATTTCCAAGAAAGCACAG
AGCTGAACAATACTTTACCTGCTTCTTACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTGAGTGAAACCCACTCACAGGCTTGTCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGACTTGTTTGTATCCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTTTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAACATCAGCATTTTAAGAAAAAAAAAA
AA

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTTCFFTISTRP
WMTQFSLLNKTCLGFFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGAACCGGAAGGCCCTGCCCAGCACTGCCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTTCTGAAGGTGCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGCTTGTCTAACCCAGGTTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSWWLATVCMLLFSHL SAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCFWLERGAGLRVTMHQPVLLCLLALIWL MVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTT
CCTGATGATTTATAGACTCAAAGAAAACTATGTTTCAGAAGCTCTCTTCTCTTCTGGCCTCCTCT
CTGTCTTCTTTCCCTCTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTATTTTTGAAATTT
CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTGTTTATGAGTATATTGCATTGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCGCGCGCCCCAACCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCGCCCCACCCACCTCCTGGCTCTTCCTGTTTTTACTCCTCTTTTCATTCTATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGAAGAATGGGGTT
CCTCGGGACCGGCACTTGGAATTCTGGTGTAGTGTCTCCGATTCAAGCTTTCCCAAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAATAAACATATCCTCCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTT
TGTTGATAAAGCTGAACTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAAGT
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
TCACTAGACGGGACTCCTTTAACCCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCAGAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCCAATAAGCCCAAGCTGGACTGAGAATCAGGCTGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGAAAGGAGAACTAAAACCTACAGTGAAGACAACCTTTGAGGAACTCCA
ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTTCAGAAAAAGCAAGAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAATATGGAACAATATCT
CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAAACAAGCT
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTGGAAGC
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGGAAATAAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCATCAATAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAATGGCAAAGATCCAGGAGTCTTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAATTTTTTGACCCAAGG
GTTATTAGAAAGTGTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC
GTAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVTMAAIQDGLAKGENDETVSNTLTTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFPAPEKSHEETDSTKEEAAMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIAKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA
GATGGTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAGCTTCTAGCTGGAGGGTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTATCCTGGGTGTCCAGGGTGGAAAG
CCAGTGCTGTCTATGTGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTCACCTTCTACCGGCGGGACATGGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCTACCCGGGCTGGTTCCGTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC
AGCAGTGTGACT**TAG**GGCAACGTGCCCCCAGAATCCCTGGGCAGAGCCAGCTCGGGTGGGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTTGTCTTCTGGTTCCAGTTTGGATAAAATCTGAGATTGGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGTACTGCTGTGGAACCTTGTA AAAACCATGTGGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCTGTCT
TAATGGTAACTGACAAGTGTACCCTGAGCCCCGAGGCCAACCCTCCCCAGTTGAGCCTTATA
GGGTGAGTACTCTCCACATGAAGTCTGTCACTCACCAGTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCATA
TGCTACCTTTCCTATCTCTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTATGTCAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTAAAAACCCAA
GATACAATCAAAATCCCAGATGCTGGTCTCTATTTCCCATGAAAAAGTGCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTAATTAAGATAACCTATTTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTAT
AGGTGATTTTCTTTTAAATCTGTTAATTTATCTGTATTTCCTAATTTTCTACAATGAAGATGA
ATTCCTTGATAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTCTGGGTAAAGGAACCTAAAGAACAAAAATCATCTGGTAATTTCTTCCTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGTGTACTGGTTGA
ATTGTGTCCCCCTCAAATTCACATCCTTCTTGGATCTCAGTCTGTGAGTTTATTTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCTATGCTGGATGAAGGTAGACCTAAATTCATAT
GACTGGTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTTGCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGGAAAGAGGCAAAGAAGAAATCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGAATAAATTTGGGCTGTTTTAA
GCCACCAAGGATAAATGGTTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAAATGATCCCTGT
CTCCTCGTGTTTACATTCTGTGTGTGTCCCTCCCAATGTACCAAAGTTGTCTTTGTGACCAA
TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTCTCTGCCACCCACCGCCCCAATCTATCTTGGCTCACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAATGAAGTCTCCT
GCCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTT
AAGTTGCTCAGTTTTTGGTCTAAGTTGTTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEIISVVPNRWLDASLSPVILGVQGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTTCCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTTACCCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTT
CAGCCTTATATGCAGGAGGTGGTGGCCCTTCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTGACACAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTTT
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFIMGTLATSCLLLLLALLVQGGAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLLIGEKLFGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCCGAGCATGTACCAGGTCACTGCAGAGGGC
TGCTTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCCTACAGGTGGTTGCAT
TCTTGGAATGGTCATGGGAACCCACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
GCCCCAACCGCCACCCAGAGTCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCACGCCCCGT
TGCCTGTGCCCCGACTGCGTCAGCCTACAGACAGGCTCCCACATGGACCCCCGGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGCGGCCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCCCTGTGAAGTGTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGAAAACCTGCACCTTCTGCACATTTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTGTCTCTCAGGAAAGGTTTTTCAA
GTTCTGCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCCTG
GCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCCTGGTTTTATT
TGTTTGTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPSKGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
GATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTTATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA
AGTGTGTCAAGGCCGAAGCCTGTGGGATCCGAACATCACTGCTTGTGAAGAAGATGAGGAGACA
GTAGAAGTGAACCTTCAACAACACTCCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCTGCTGTCTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAAACCATTGCAGAAGTGAGGTCATCCTTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCA
GCAGACAAAGTCGTCTTCTTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTCTCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLRDLRVEPVTTSVATGDYSILMNVSWV
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMYKKKCVKAGSLWDPNITACKNEETVEVNFTTTPLGNYMALIQH
STIIGFSQVFEPHQKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
SKPGGWLPLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGPSSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKKQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTTCCAAAAGCCTGAGAGTTGCCCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 16o

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTWTWDPNRYPSSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGC SVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
ACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGGATTTC
CAGCCCCCTGCCACCCACAGACACGGGTGACTGGGGTGTCTGCCCCCTTGGGGGGGGGCAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTTCTTGCTGTCTT
TGGCACTGGGCCGAAGCCAGTGGTCTTTCTCTGGAGAGGCTTGTGGGGCCTCAGGACGCTACC
CACTGCTCTCCGGGCTCTCCTGCCGCTCTGGGACAGTGACATACTCTGCCCTGCCCTGGGGACAT
CGTGCCCTGCTCCGGGCCCCGTGCTGGCGCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCCGTGCATGGGCACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
TGCCTCTCTCCAGGCCCAAGTCGTGCTCTCCTTCCAGGCCTACCTACTGCCCGCTGCGTCTGCTG
TGGAGGTGCAAGTGCCCTGCTGCCCTTGTGCAGTTTGGTCAGTCTGTGGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCTTATACTCAGCCCAGGTACGAGAA
GGAACCTCAACACACACAGCAGCTGCCCTGCCCTGCCCTGGCTCAACGTGTGAGCAGATGGTGACA
ACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCACTTCGGCCTCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCCAAAACCCCGGTGGCACAAAAACCTGACTGGACCGCAGATCATTACCTTGAA
CCACACAGACCTGGTTCCCTGCCTCTGTATTTCAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA
CGAATCTGCCCCCTTCAAGGAGGACCCCCGCGCACACCAGAACCTCTGGCAAGCCGCCGCTG
CGACTGTGACCTGCAGAGCTGGCTGCTGGACGACCGTGTGCTGCTGCCCGCAGAAGCGGCACT
GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCACTGGTCCCACCGCTTTCCTGGGAGAACG
TCACTGTGGACAAGTTCTCGAGTTCCCATTTGCTGAAAGGCCACCCTAACCTCTGTGTTTCAAGTG
AACAGCTCGGAGAAGCTGCAGCTGCAGGAGTGCTTGTGGGCTGACTCCCTGGGGCCTCTCAAAGA
CGATGTGCTACTGTGGAGACACGAGGCCCCCAGGACAACAGATCCCTCTGTGCCTTGGAACCCA
GTGGCTGTACTTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCTG
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCTACTCTTTGCCG
CTGCGCTTTCCCTCATCTCTCTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAA
CAGGACGTCCGCTCGGGGGCGGCCGCGAGGGCCGCGCGGCTGTGCTCCTTACTCAGCCGATGA
CTCGGGTTTCGAGCGCTGGTGGGCGCCCTGGCGTCCGCGCTGTGCCAGCTGCCGCTGCGCGTGG
CCGTAGACCTGTGGAGCCGTCTGTAAGTGAAGCGCGCAGGGGCCCGTGGCTTGGTTTACGCGCAG
CGGCGCCAGACCCTGCAGGAGGGCGCGTGGTGGTCTTGCTCTTCTCTCCCGGTGCGGTGGCGCT
GTGCAGCGAGTGGCTACAGGATGGGGTGTCCGGGCCCCGGGCGCACGGCCCCGACGACGCTTCC
GCGCCTCGCTCAGCTGCGTGTGCCGACTTCTTGACAGGGCCGGGCGCCCGGAGCTACGTGGGG
GCCTGCTTCGACAGGCTGCTCCACCCGACGCCGTACCCGCCCTTTTCCGCACCGTGCCCGTCTT
CACACTGCCCTCCCAACTGCCAGACTTCTTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCGTCCG
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTCAAGCCAGCCCTGGATAGCTACTTC
CATCCCCGGGGACTCCCCGCGCGGGACGCGGGGTGGGACCAGGGGCGGGACCTGGGGCGGGGA
CGGGACTTAAATAAAGGCAGACGCTGTTTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSELRVGPQDATHCSPGLSCRLWDSILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSTFQAYPTARCVLLEV
QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSDAGDNVHLVLNV
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQ
LWQAARLRLTLQSWLLDAPCSLPAAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQ
VNSSEKLQEQECLWADSLGPLKDDVLLLETRGPQDNRSICALEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLLLLKKDHAKGWLRLKQDVRSGAAARG
RAALLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFWHAQRRQTLQEGGVVLLFSP
GAVALCSEWLQDGVSGFGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT
LPSQLPDFLGALQQPRAPRSGRLQERAEQVSRAALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
 GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA
 AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
 AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTGACGGGATCACCCGGAAGTCTGC
 AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
 GTCAGTGCGGGAGGCGGTCAGCCACCAAGATGACTGACAGGTTGAGTCTCTGCAGCACACTAC
 CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGATTGTTTCACCTA
 CCCCCACGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCCGGAAGACATCTTCCATGACCTG
 TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAGCAGAGAGA
 ATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTCCTTGGCACCATCATGATTGCGTTCCCA
 CCTGGGCCAAGGAGAGTGGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
 TACTCCTTCTCCGGAGCCTTCTGTTCTCCATGGGCTTCTCGTCGAGTACTCTGCTACCTGAG
 CTACAGATATGTACCAAGCGCCTGCACCTCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
 TCCAGCGCTGCGCTTCATCCAGGAGCACGTCCCTGATCCCTGTCTTTGACCTCAGCGGCCCAGC
 AGTCTGGCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
 TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
 CCTCCACGTGCCACCTCCCCAGATCCTCTCCCACTGTCTATGCCCCAAACGCTGCCCTGAG
 GTCGGGCCCCCATCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCCATTTACGCCCCACA
 GGCCATCTCTAAGTCCAGCCTTCTCCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
 CCTATGGGGTATGCATGGAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCTTAA
 CACCTTAGGCCTAAAGGTGAGCTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
 TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATTGCACCAGC
 CCCTGGGGATTTGCACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGCACA
 CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
 TGCTGGAGTCCCTTGTTGTGTCCTCAAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
 GAGCAGCCCCACAGAAGTGGATTCTCTTTTCAGAGGCTGGCCCTGACTGTGAGTGGGAGTCCCTG
 AGGGGAATGGGAAAGGCTTGGTGCTTCTCCTGTCCTACCCAGTGTACATCCTTGGCTGTCA
 ATCCCATGCCTGCCCATGCCACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
 AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTACCGGAACAAAGCAGCATGATA
 AGGACTGCAGCGGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCTGCTCGCTGAGCCCTG
 CAAGGCAGAAATGACAGTGAAGGAGGAAATGCAGGGAAATCCCGAGGTCCAGAGCCCCACCTC
 CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATTCTTGCCAGTTT
 ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCTCTTCTGTCATTGTTCAAAGGTGGGAAGAGA
 GCCTGGAAAAGAACCAGGCTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAGAACAACCTGC
 ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCA
 TTCCCAGCCAGGGCACTGCCTGACGTTGCACGATTTAGCTTCATTCTCTGATAGAACAAAGC
 GAAATGCAGGTCCACCAGGGAGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTACAGCCCT
 ATCCTGAGAATGGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
 TGTACTGATGTACAACTTTGCAAGCTCTGCCTTGGGTTTCAGCCCATCTGGGCTCAAATTCAGC
 CTCACCACTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTTCTC
 ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
 TCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGCCCAATAAACGGTAGCTATTTAAAAA
 AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTelyarvtavsaggrsatkmtDRFSSLQHTTLKPPDVCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCrvkTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLTfQPLRFIQEHVLIpVFDLSGPSSLAQPvQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPQILSPLSYAPNAapevgPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSLAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEETPQYLKGQLPLLSSVQIEGHMPSLPLQPPSGPC
SPSDQGPSPWGLLESVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTvQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAGTCACCCGGGCGCGGTGGCCACAACATGG
CTGCGGCGCGGGGGCTGCTCTTCTGGCTGTTCTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACTTTGGGCTGGA
AGTGTGGAACACAGTTTTTGGATATTTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCCTGTGCCTGAGCCCCGAGGCATTTCAGAGCTGATTTCAGAGGATGGAGAAGGTGCTTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTTGTTGGCATCAAAATCTAAGTTTGTTTTACAAAGATTGTTTTTAGTA
CTAAGCTGCCTTGGCAGTTTGCATTTTGTAGCCAAACAAAAATATATTATTTTCCCTTCTAAGTA
AAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFWLGALWWVPGQSDLSHGRRFSDLKVCGDEECSSMLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPADETDFVCFEGGRD
DFNSYNVEELLGSLELEDVPEESKKAEEVSQHREKSPESRGRELDPVPEPEAFRADSEDGEGA
FSESTEGLOGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGESERTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCTCAGAGGCCGGGGAAGAGAAGCAAAGCGC
 AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC
 AGTCCCCCAAACGCGCACCTCGAAGTCTTGAATCCAGCCCCGCACATCCACGCGCGGCACAGG
 CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGTCGGGCAGCTGGGCTCGGGC
 GCGGGAGTAGGGCCCCGCGAGGGAGGCGAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
 GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
 TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCGCGCCGCTGGTCAGC
 GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAACT
 GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCCTGGCTTGTGAGAGTGAGGGAGGAGTCTCTCTCA
 GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
 ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGG
 TGCCTGCCCAGATCTTACCAGTGGTCTGATGGAAGCAATCCCAAGTACCGAAACTGGTACACAG
 ATGAACCTTCTTGGGAAGTGAAAAGTGTGTTGTGATGTATACCAACCAACTGCCAATCTCTGGC
 CTTGGGGGTCCTTACCTTTACCAGTGGAAATGATGACAGGTGTAACATGAAGCACAATTTATTTTG
 CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG
 GAGACACCCATCAGAATGTGGTGTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT
 ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGAACCTGTGTTTCCAGATGCT
 GCATAAAAGTAAAGGAAGAACAAAACCTAGTCCAAACAGTCTACACTGTGGATTTCAAGAGTA
 CCAGAAAAGAAAGTGGCATGGAAGTATAAATACTCATTGACTTGGTTCCAGAATTTTGAATTTCT
 GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGAATGGCTTGAATCACAAGGATCTGC
 AAGATGAAGTGAAGCTCCCCCTTGAGGCAAAATATTAAAGTAATTTTATATGCTATTATTTCA
 TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGTCTAAAGGATGCACCCAA
 ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
 TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTCACCTTTCATAAGTTGTTATCTAGTCAATGTAA
 TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGTTGATAAAA
 ATGAAGTGTCTAATATTTATTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAG
 AAACCTTATTACTGTTGTCAACTGAATTCACACACACACAAAATATAGTACCATAGAAAAAGTTTGT
 TTTCTCGAAATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
 AATAAGAAGCTATTTTCAATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
 TGTCTAATTTCAATGTGCAAGACATGTGCCTTATAATTATTTTAGCTTAAATTAACAGATT
 TTGTAATAATGTAACCTTGTAAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAG
 TGACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA
 GGGTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTTCGGGGTTTGGG
 ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
 ACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
 ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAAATGGGTTGG
 AACCCATCAGTGATCGCATATTCTTATGATGAGGGTTTGCTTGAGATAGAAAATGGTGGCTCCTTT
 CTGCTTATCTCTAGTTTCTTCAATGCTTACGCCTTGTCTTCTCAAGAGAAAGTTGTAACCTCT
 CTGGTCTCATATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTTCTTGTTCTGGGGGAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFCRRVVSQGKVCFADEFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSCKGRKTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217